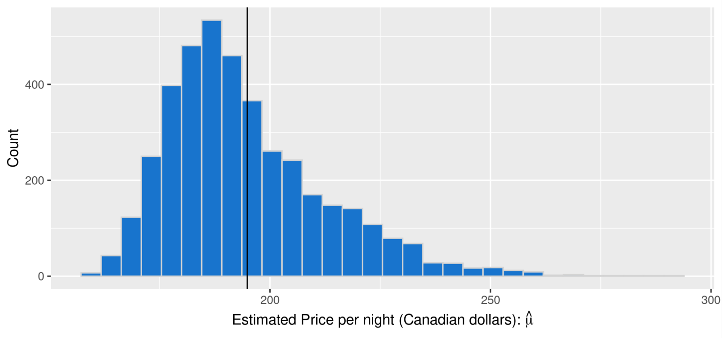
Dataset

Recall our running example from previous classes:[1]

- We're interested in the mean price of AirBNBs in our city (μ)
- We can't observe them all, so we take the mean price of a sample of 200 ($\hat{\mu}(X) = \frac{1}{200} \sum_{n=1}^{200} x_n$)

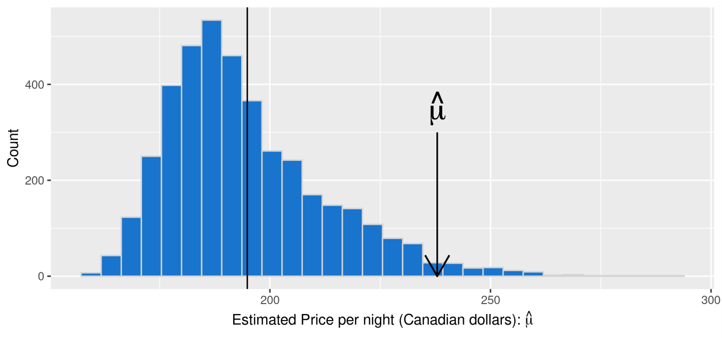


[1] Taken from Chapter 10 of "Data Science: A First Introduction" by Timbers, Campbell, and Lee https://ubc-dsci.github.io/introduction-to-datascience/

Dataset

Recall our running example from previous classes:[1]

- We're interested in the mean price of AirBNBs in our city (μ)
- We can't observe them all, so we take the mean price of a sample of 200 ($\hat{\mu}(X) = \frac{1}{200}\sum_{n=1}^{200} x_n$)

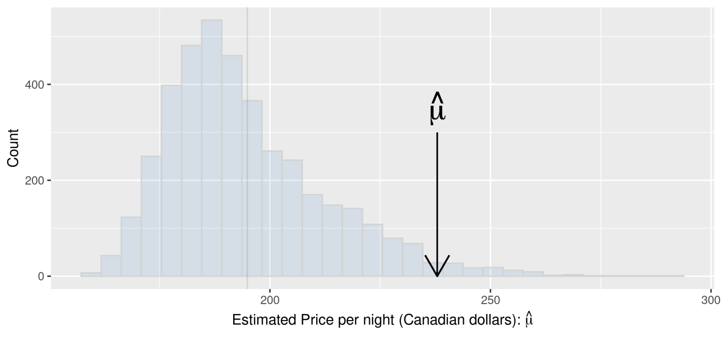


[1] Taken from Chapter 10 of "Data Science: A First Introduction" by Timbers, Campbell, and Lee `https://ubc-dsci.github.io/introduction-to-datascience/`

Dataset

Recall our running example from previous classes:[1]

- We're interested in the mean price of AirBNBs in our city (μ)
- We can't observe them all, so we take the mean price of a sample of 200 ($\hat{\mu}(X) = \frac{1}{200}\sum_{n=1}^{200} x_n$)



[1] Taken from Chapter 10 of "Data Science: A First Introduction" by Timbers, Campbell, and Lee https://ubc-dsci.github.io/introduction-to-datascience/

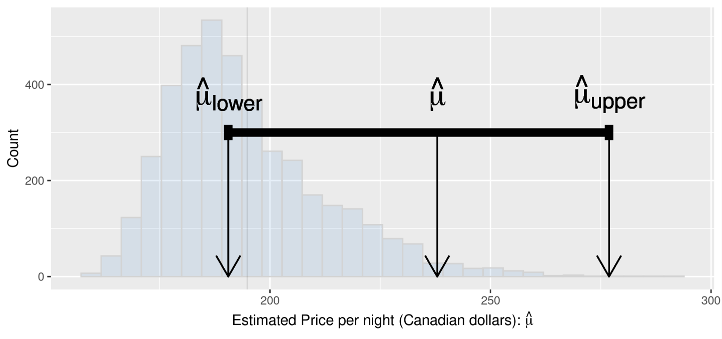
Dataset

Recall our running example from previous classes:[1]

- We're interested in the mean price of AirBNBs in our city (μ)
- We can't observe them all, so we take the mean price of a sample of 200 ($\hat{\mu}(X) = \frac{1}{200} \sum_{n=1}^{200} x_n$)

Key idea:

Instead of a "point estimate" $\hat{\mu}(X)$, estimate an interval $(\hat{\mu}_{lower}(X), \hat{\mu}_{upper}(X))$.

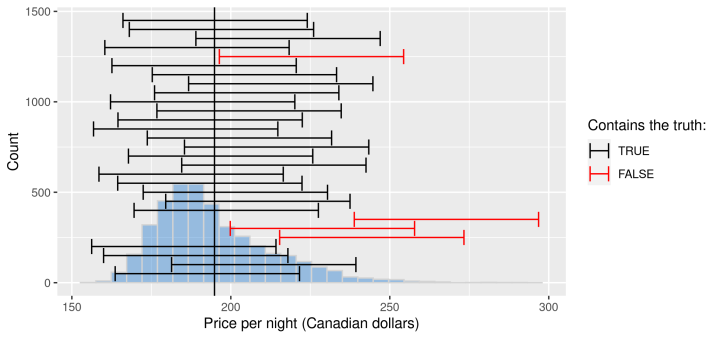


[1] Taken from Chapter 10 of "Data Science: A First Introduction" by Timbers, Campbell, and Lee https://ubc-dsci.github.io/introduction-to-datascience/

Dataset

Key idea:

Instead of a "point estimate" $\hat{\mu}(X)$, estimate an interval $(\hat{\mu}_{lower}(X), \hat{\mu}_{upper}(X))$.



Dataset

Key idea:

Instead of a "point estimate" $\hat{\mu}(X)$, estimate an interval $(\hat{\mu}_{lower}(X), \hat{\mu}_{upper}(X))$.

We would like to choose our interval such that:

$P(\mu \in (\hat{\mu}_{lower}(X), \hat{\mu}_{upper}(X))) \geq 0.9$

Such an interval is a **valid confidence interval** with a **level** of 0.9.